

Research Paper

High-content stimulated Raman histology of human breast cancer

Hongli Ni^{1,#}, Chinmayee Prabhu Dessai^{2,#}, Haonan Lin^{1,#}, Wei Wang³, Shaoxiong Chen⁴, Yuhao Yuan¹, Xiaowei Ge¹, Jianpeng Ao¹, Nolan Vild¹, and Ji-Xin Cheng^{1,2}✉

1. Department of Electrical and Computer Engineering, Boston University, 8 St. Mary's St., Boston, MA, 02215, USA.
2. Department of Biomedical Engineering, Boston University, 44 Cummington Mall, MA 02215, USA.
3. Hologic Inc., 250 campus drive, Marlborough, MA 01752, USA.
4. Indiana University School of Medicine 340 West 10th Street, Fairbanks Hall, Suite 6200, IN 46202, USA.

#These authors contributed equally.

✉ Corresponding author: jxcheng@bu.edu.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2023.09.20; Accepted: 2023.12.17; Published: 2024.01.27

Abstract

Histological examination is crucial for cancer diagnosis, however, the labor-intensive sample preparation involved in the histology impedes the speed of diagnosis. Recently developed two-color stimulated Raman histology could bypass the complex tissue processing to generates result close to hematoxylin and eosin staining, which is one of the golden standards in cancer histology. Yet, the underlying chemical features are not revealed in two-color stimulated Raman histology, compromising the effectiveness of prognostic stratification. Here, we present a high-content stimulated Raman histology (HC-SRH) platform that provides both morphological and chemical information for cancer diagnosis based on un-stained breast tissues.

Methods: By utilizing both hyperspectral SRS imaging in the C-H vibration window and sparsity-penalized unmixing of overlapped spectral profiles, HC-SRH enabled high-content chemical mapping of saturated lipids, unsaturated lipids, cellular protein, extracellular matrix (ECM), and water. Spectral selective sampling was further implemented to boost the speed of HC-SRH. To show the potential for clinical use, HC-SRH using a compact fiber laser-based stimulated Raman microscope was demonstrated. Harnessing the wide and rapid tuning capability of the fiber laser, both C-H and fingerprint vibration windows were accessed.

Results: HC-SRH successfully mapped unsaturated lipids, cellular protein, extracellular matrix, saturated lipid, and water in breast tissue. With these five chemical maps, HC-SRH provided distinct contrast for tissue components including duct, stroma, fat cell, necrosis, and vessel. With selective spectral sampling, the speed of HC-SRH was improved by one order of magnitude. The fiber-laser-based HC-SRH produced the same image quality in the C-H window as the state-of-the-art solid laser. In the fingerprint window, nucleic acid and solid-state ester contrast was demonstrated.

Conclusions: HC-SRH provides both morphological and chemical information of tissue in a label-free manner. The chemical information detected is beyond the reach of traditional hematoxylin and eosin staining and heralds the potential of HC-SRH for biomarker discovery.

Keywords: stimulated Raman histology, high-content chemical imaging, breast cancer, label-free histology, fiber laser

Introduction

Breast cancer is one of the most prevalent cancers, accounting for ~15% of cancer-related deaths in women [1, 2]. As a highly heterogeneous disease, breast cancer requires detailed histopathology for

diagnosis and treatment. However, the staining-based histology requires labor-intensive and time-consuming sample preparation, which inhibits rapid diagnosis. Besides, the inter- and intra-lab differences

in the staining protocol can induce result variation, leading to a diagnosis discordance [3, 4]. Staining differentiates tissue components based on their chemical contents, so tissue imaging techniques that require minimum sample preparation while providing rich chemical information are good candidates to overcome the above-mentioned problems.

As a label-free non-destructive chemical analysis approach, Raman spectroscopy has been widely applied in cancer research [5, 6]. Spontaneous Raman spectroscopy can hardly be applied to tissue histology due to its slow signal acquisition speed, typically several seconds per spectrum. Coherent Raman scattering microscopy, based on either coherent anti-Stokes Raman scattering (CARS) or stimulated Raman scattering (SRS), overcomes this limitation through coherent excitation, enabling high-speed chemical imaging [7-9]. Compared to CARS, SRS is more favorable for clinical practice because SRS is free of non-resonant background and can be performed under ambient light. SRS microscopy has been used for label-free cancer histology. By detecting lipid-rich and protein-rich regions with two SRS spectral channels, two-color stimulated Raman histology (SRH) was developed to visualize brain tumor morphology [10, 11]. The two-color SRH results correlate well with the hematoxylin and eosin staining (H&E) data, which enables rapid intraoperative brain tumor analysis when combined with deep learning algorithms [12, 13]. The two-color SRH was also successfully applied to other cancer types, including gastric, prostate, and breast cancer [14-16]. Besides two-color SRH, other studies utilized SRS to map microcalcifications which is a well-known breast lesion biomarker, for breast tissue classification [17, 18]. The two-color SRH and microcalcification-based approach both achieved high breast cancer detection accuracy when combined with machine learning models. Nevertheless, these methods only detect one or two chemical components, which compromises the chemical mapping capability of SRS. Recent studies show that spectroscopic SRS is capable of targeting multiple biomolecules of interest [19, 20]. As cancer development involves a variety of biomolecules, providing such rich chemical information is important for accurate subtyping of breast cancer [21, 22] and enables precise treatment.

Here, we demonstrated high-content stimulated Raman histology (HC-SRH) that exploits the rich chemical information in spectroscopic SRS. By implementing hyperspectral SRS imaging in the C-H spectral window and subsequent sparsity-penalized unmixing, we mapped 5 major chemical components in breast tissue: saturated lipids, unsaturated lipids,

cellular protein, extracellular matrix (ECM), and water. The change in amount and relative ratio of these components were shown to be closely related to cancer progression [23-26]. In our work, successful high-content chemical mapping in the crowded C-H window was achieved by a least absolute shrinkage and selection operator (LASSO) regression algorithm for spectral unmixing [27, 28]. As imaging speed is a key parameter for clinical trials, we further boosted the speed of HC-SRH by one order of magnitude through spectrally selective sampling. Our HC-SRH results showed various tissue components for sub-typing of breast cancers. Moreover, using a multi-window fiber laser-based SRS microscope [29], we demonstrate compact HC-SRH for which the results agree well with the state-of-the-art SRS system. In the meanwhile, the broader spectral range of the multi-window fiber laser enabled HC-SRH to map more types of biomolecules. Our fingerprint results showcased a clear contrast for nucleic acid and solid-state ester as well as representative peaks for different amino acids. These contrasts are beyond the reach of C-H based SRH and show the potential of fingerprint SRH for discovering new biomarkers for cancer diagnosis [22].

Methods and materials

SRS imaging systems and data acquisition

Figure 1A shows the setup of a state-of-the-art SRS system based on a solid-state optical parametric oscillator (OPO). Without extra note, the SRS data was collected with this system. The solid-state OPO outputs two synchronized femtosecond laser pulse trains with 80 MHz repetition rate. The laser with shorter wavelength provides the pump beam and the other one fixed at 1040 nm provides the Stokes beam. For HC-SRH, the pump wavelength is centered at 800 nm. The Stokes beam is modulated by an acousto-optic modulator (AOM) at 2.4 MHz and pre-chirped by 15 cm SF57 rods. A motorized delay stage is put in the pump path to control the temporal overlapping of the two laser beams. The pump and Stokes are spatially combined with a dichroic mirror and then linearly chirped by 75 cm long SF57 glass rods. The combined beam is focused on the sample by a 60X water-immersion objective with 1.2 numerical aperture (NA). The transmission light is collected by an oil-immersion condenser with 1.4 NA and then filtered by a 1000 nm short pass filter to block the Stokes beam. The transmitted pump beam is detected by a photodiode. The SRS signal is conveyed as the modulation transfer from the Stokes beam to the pump beam. Extraction of the modulation is done with a lock-in amplifier. An SRS image is formed by

laser scanning with a galvo mirror. A motorized sample stage is combined with the galvo mirror to cover a large field-of-view (FOV) for large-area tissue mapping. Hyperspectral SRS imaging is achieved through spectral focusing where the targeted spectral position is controlled by the temporal delay between the linearly chirped pump and Stokes laser pulses [19]. Second harmonic generation (SHG) imaging is performed in the same setup by replacing the photodiode with a photomultiplier tube. When performing SHG, only the Stokes laser is incident on the sample and a 520/70 bandpass filter is placed before the photomultiplier tube to filter out the Stokes beam.

Figure 1B presents the schematic of a compact SRS system based on a fiber OPO (FOPO) [30]. The setup of the FOPO-based SRS system is the same as the solid-state OPO-based system except for the modulation, hyperspectral scanning, and detection parts. The fiber laser outputs two synchronized picosecond laser pulse trains with 40 MHz repetition rate. The Stokes is modulated by a built-in modulator inside the laser at 6.7 MHz. This system realizes hyperspectral SRS by tuning the laser wavelength; therefore, no chirping rods and motorized stage are required. The FOPO can achieve arbitrary tuning between 700 and 3100 cm^{-1} within 20 milliseconds. Because the fiber laser has high laser intensity noise, auto-balanced detection is implemented to improve signal-to-noise ratio. Briefly, two identical photodiodes detect the transmitted pump beam and a reference pump beam without passing the sample.

The signal amplitudes from the two photodiodes are matched automatically by the variable gain amplifier and the proportional-integral-derivative controller. The laser noise is then suppressed by performing difference of the two amplitude-matched signals. The SRS signal is obtained by demodulating the differential signal with a lock-in amplifier.

The laser power incident on the sample is set as follows: pump 15 mW, Stokes 50 mW for solid-state OPO-based SRS; Stokes 50 mW for SHG; pump 15 mW, Stokes 100 mW for fiber laser-based SRS. All the SRS data is acquired with 10 μs pixel dwell time and 500 nm pixel size. The imaging time is around 1.0 minute for 1.0X1.0 mm^2 per spectral channel.

Sample preparation

The human breast cancer tissue samples are freshly frozen tissue sections purchased from Biochain Institute (Newark, CA). The tissue samples were collected by the company from anonymous patients with ethics approval. Each tissue section has a thickness around 5~10 μm and was mounted on a standard glass slide. For SRS and SHG imaging, the frozen tissue slide was washed using 1X phosphate-buffered saline (PBS) and then covered with a glass coverslip. Nail polish was used to seal the coverslip to prevent tissue dehydration. To acquire histology, a neighboring section of the tissue section used for SRS was prepared for standard H&E staining. The co-registration of HC-SRH and H&E was done manually.

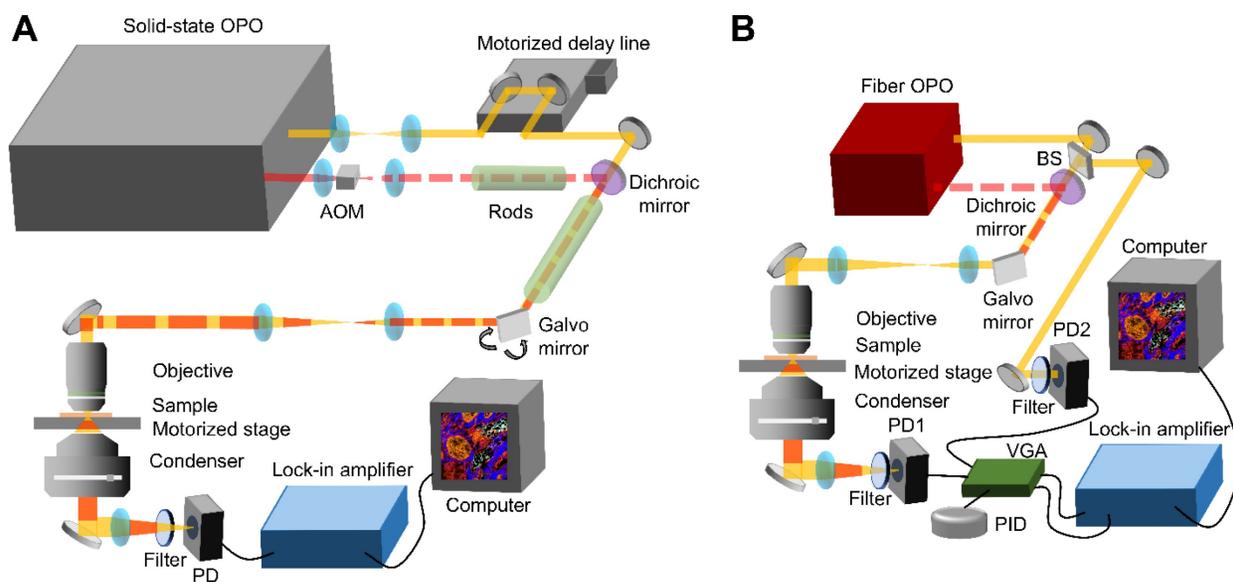


Figure 1. System schematic. (A) A state-of-the-art SRS imaging system with a solid-state laser. (B) A compact fiber laser-based SRS imaging system. OPO: optical parametric oscillator. AOM: acousto-optic modulator. PD: Photodiode. VGA: Variable gain amplifier. PID: proportional-integral-derivative controller. BS: Beam Splitter

Data processing

For every large-area SRS imaging data, the small-FOV SRS hyperspectral images were first fused into a large-FOV image with the ImageJ grid stitching toolbox. The fused hyperspectral image was then denoised by the spectral total variation algorithm [31]. Finally, the denoised SRS data was fed into the LASSO algorithm with reference spectra for generating concentration maps. The principle of LASSO can be expressed as:

$$\hat{C}_i = \underset{C_i}{\operatorname{argmin}} \left\{ \frac{1}{2} \|D(i, :) - C_i S\|^2 + \beta \|C_i\|_1 \right\}$$

Where $D(i, :)$ is the spectrum of the i^{th} pixel in the data, C_i represents the concentration for the targeted chemical components, S is the reference spectra for the targeted chemical components. β is the hyperparameter that can be fine-tuned to minimize channel crosstalk. We fixed $\beta = 0.01$ in this work. The reference spectra for unsaturated lipids, ECM, saturated lipids, and water were acquired from glycerol trioleate, breast tissue ECM, palmitic acid, and 1X PBS, respectively. The cellular protein spectrum is extracted by subtracting the unsaturated lipid spectrum from the breast cancer cell spectrum.

Spectral selective sampling is performed by a recursive feature elimination (RFE) algorithm. The principle of RFE is to iteratively discard the least contributing channel for the LASSO unmixing result. To determine the least important channel, every channel in the current spectrum will be tentatively discarded and a mean square error (MSE) will be calculated between the current LASSO result and the LASSO result generated with the full spectrum. The channel corresponding to the minimum MSE is the least important. The RFE algorithm stops when the number of current spectral channels reaches the targeted value. Because RFE has high computational complexity, we only fed a small data set of 2100 pixels to the algorithm to accelerate the selection process. The 2100 pixels consist of 5 subsets, where every subset has a dominant chemical component and was randomly selected from a representative FOV. This design can balance the contribution of every chemical component in the selection process. The data processing flow for the spectral selective sampled SRS data is the same as the hyperspectral data.

Results

HC-SRH maps 5 chemical contents of breast cancer tissue

Figure 2 demonstrates that HC-SRH in the C-H window can successfully map 5 major chemical components in breast cancer tissue. **Figure 2A** shows

the merged pseudo-color concentration map with yellow for unsaturated lipids, red for cellular protein, blue for ECM, cyan for saturated lipids, and grey for water. The separated concentration maps are shown in **Figure 2B**, where the water channel is contrast-enhanced to visualize the details better. The reference spectra used for LASSO unmixing are also shown in **Figure 2B**. The major spectral difference between the cellular protein and ECM is the overall spectral blue shift. This blue shift can come from the increased hydrogen bonds [32] or can be related to the fibrils formation [33].

To validate the concentration mapping given by HC-SRH, we compared the results of HC-SRH imaging with standard H&E staining on adjacent tissue section (**Figure 2C**). The unsaturated lipid is rich in the cell cytoplasm and fat cell residues, which is corroborated by the literature that breast cancer cells and adipose cells contain rich unsaturated lipids [34, 35]. The fat cells are seen as holes in HC-SRH because the tissue slicing procedure punctured the fat cells and only membrane-like residues remained. The cellular protein and ECM maps in HC-SRH agree well with the cell and ECM area in the H&E result. The fat cell residue area also shows a high saturated fat content in HC-SRH, which is consistent with that breast fat contains rich saturated lipids [34]. In addition, the solid-form saturated lipids are more likely to be preserved in tissue slicing. The water channel indicates that water is distributed across the whole FOV while the ECM contains slightly higher water content, and the fat residue area is lower in water content. This distribution makes sense because the main constituent of ECM, collagen, is hydrophilic, whereas lipid is hydrophobic. An enlarged view of one small FOV corresponding to the white box in **Figure 2A** is shown in **Figure 2D** with spectra of five representative pixels for cell nuclei, cell cytoplasm, ECM, fat cell residue and empty area.

We further compared HC-SRH with SHG which is commonly used for collagen imaging (**Figure 2E**). A cancer-adjacent blood vessel FOV is selected for verification because this FOV is rich in ECM including collagen and elastin. The H&E result is acquired from an adjacent section from the section used for SRS and SHG. The ECM channel of the HC-SRH correlates well with the SHG result. Yet, there are two main differences between the ECM channel of HC-SRH and the SHG. First, the intensity differs significantly for large and small fibers in SHG but appears more uniform in HC-SRH ECM channel. This difference may be because SHG has a strong signal for type I and II collagen as well as highly aligned fibers but less sensitive to other types of collagens and small fibers [36]. The SRS signal is based on the protein

concentration therefore provides better contrast for the small fibers. The second difference is that the blood vessel wall is visible in HC-SRH ECM channel but not in SHG. This difference is because the ECM of the blood vessel wall is rich in elastin [37], which can be detected by SRS but has a very low SHG signal.

Selective spectral sampling boosts HC-SRH speed by one order of magnitude

The HC-SRH results shown in **Figure 2A** are generated with 45 SRS spectral channels covering 2820 to 3030 cm^{-1} , which required 40 minutes for the

full field of view stack. As imaging speed is crucial for clinical practice, we boosted the speed of HC-SRH by reducing the number of spectral channels to be sampled. Theoretically, 5 chemical components require at least 5 channels to unmix. Therefore, we implemented the RFE algorithm to reduce the spectral channels to 5 while monitoring the HC-SRH result degradation to find a balance of between speed and image quality. **Figure 3** presents the HC-SRH result of two breast cancer tissue sections with 45, 20, 10, and 5 selected spectral channels. The HC-SRH quality is quantified by the structural similarity (SSIM) and

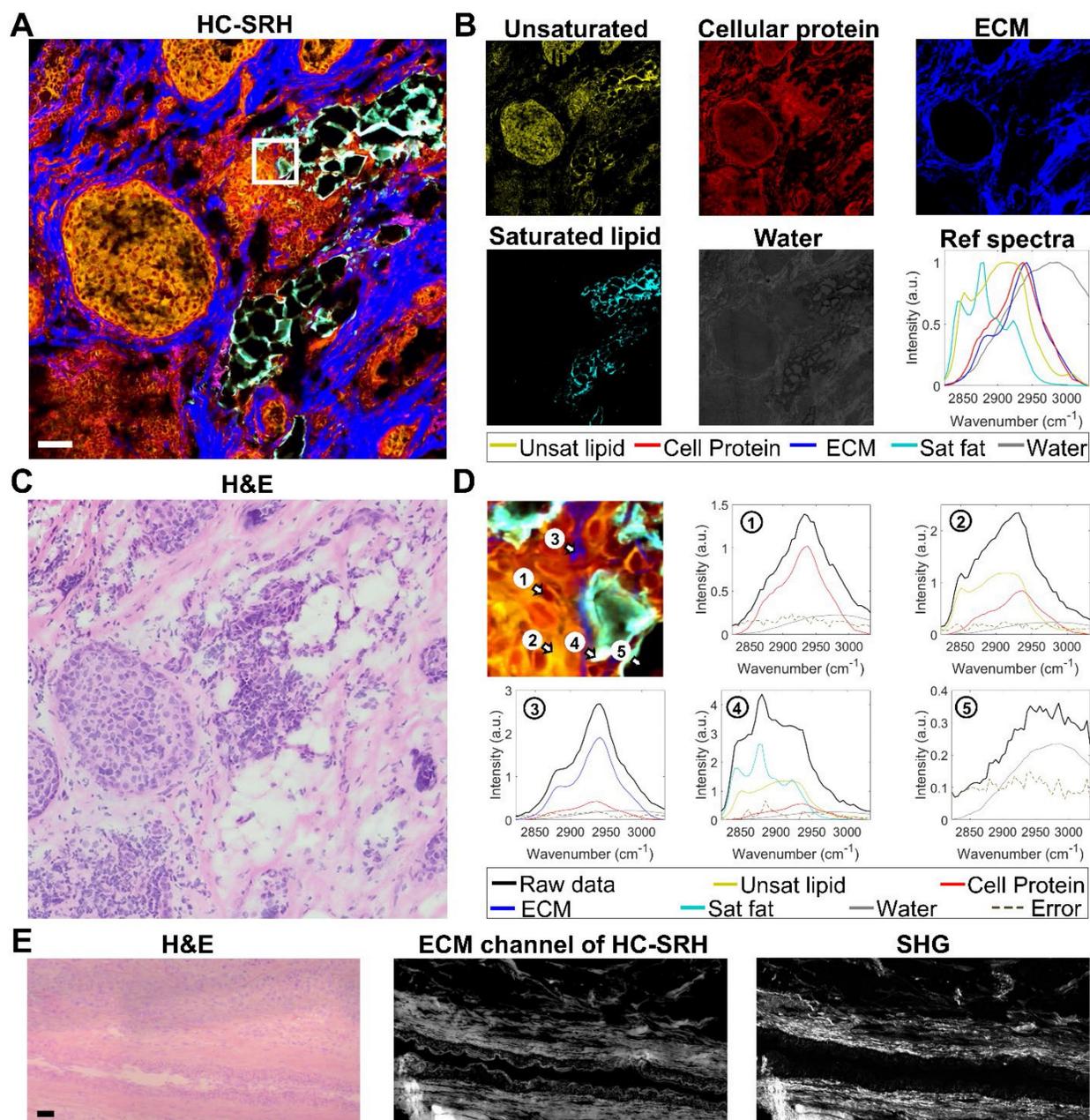


Figure 2. HC-SRH of human breast cancer tissue slice. (A) Merged concentration map for a breast cancer tissue sample. Yellow for unsaturated lipids, red for cellular protein, blue for ECM, cyan for saturated lipids, and grey for water. (B) Separate concentration maps for the 5 chemical components and the reference spectra corresponding to (A). (C) H&E result of a neighboring section to the tissue section used in (A). (D) Zoom-in view of the white box in (a) and SRS spectra of the five selected pixels as numbered. (E) Comparison of HC-SRH and SHG in mapping ECM in a cancer-adjacent vessel. Scale bar: 50 μm .

peak signal-to-noise ratio (PSNR) to the 45-channel result. The channel reduction does not lead to significant image quality degradation. Even with only 5 channels, the image quality is sufficiently enough for further analysis. Hence, the speed of HC-SRH can be boosted by 9 times by reducing 45 spectral channels to 5 selected ones (2837, 2880, 2908, 2951, and 3009 cm^{-1}). When using different datasets for the RFE selection, the selected channels were slightly different but proved to have similar performance (Supplementary material section 1). Thus, there can be multiple good combinations for the selective sampling, which makes sense as the reference spectra in the C-H window cover a wide spectral range.

HC-SRH provides excellent contrast for various breast tissue components

With the 5 selected spectral channels, we demonstrated the histology capability of our method by imaging different types of breast lesion and structure, including typical ductal carcinoma in-situ (DCIS), usual ductal hyperplasia (UDH), normal duct, invasive ductal carcinoma (IDC), invasive lobular carcinoma (ILC), and chemotherapy lesion (Figure

4A-D). Figure 4E-H shows the H&E results of adjacent sections corresponding to Figure 4A-D. The SRH acquisition time for every tissue is less than 7 minutes.

We selected 10 regions of interest (ROI) from the four breast tissue sections to illustrate the performance of HC-SRH in detail. The ROI 1 is a typical DCIS where the cancer cells have grown inside the duct, and the ECM fibers nearby are stretched because of the over-growth. ROI 2 shows a necrosis area in the center of the DCIS duct, which is a signature of high-grade DCIS [38]. In HC-SRH, the necrosis shows an obviously higher cellular protein concentration than the surrounding cells, possibly due to cell condensation [39]. ROI 3 visualizes some stroma cells buried in the ECM, whose zoom-in view is shown in Figure 4I. ROI 4 shows IDC permeates the fat cell area, where the fat cell morphology differs from the healthy fat cells with smaller structures surrounded by more ECM, which is the signature of cancer-associated adipocytes. (Figure 4J). ROI 5 presents a normal duct structure. ROI 6 is a usual ductal hyperplasia structure whose zoom-in is Figure 4K. The duct cells in Figure 4K are overgrown, but the

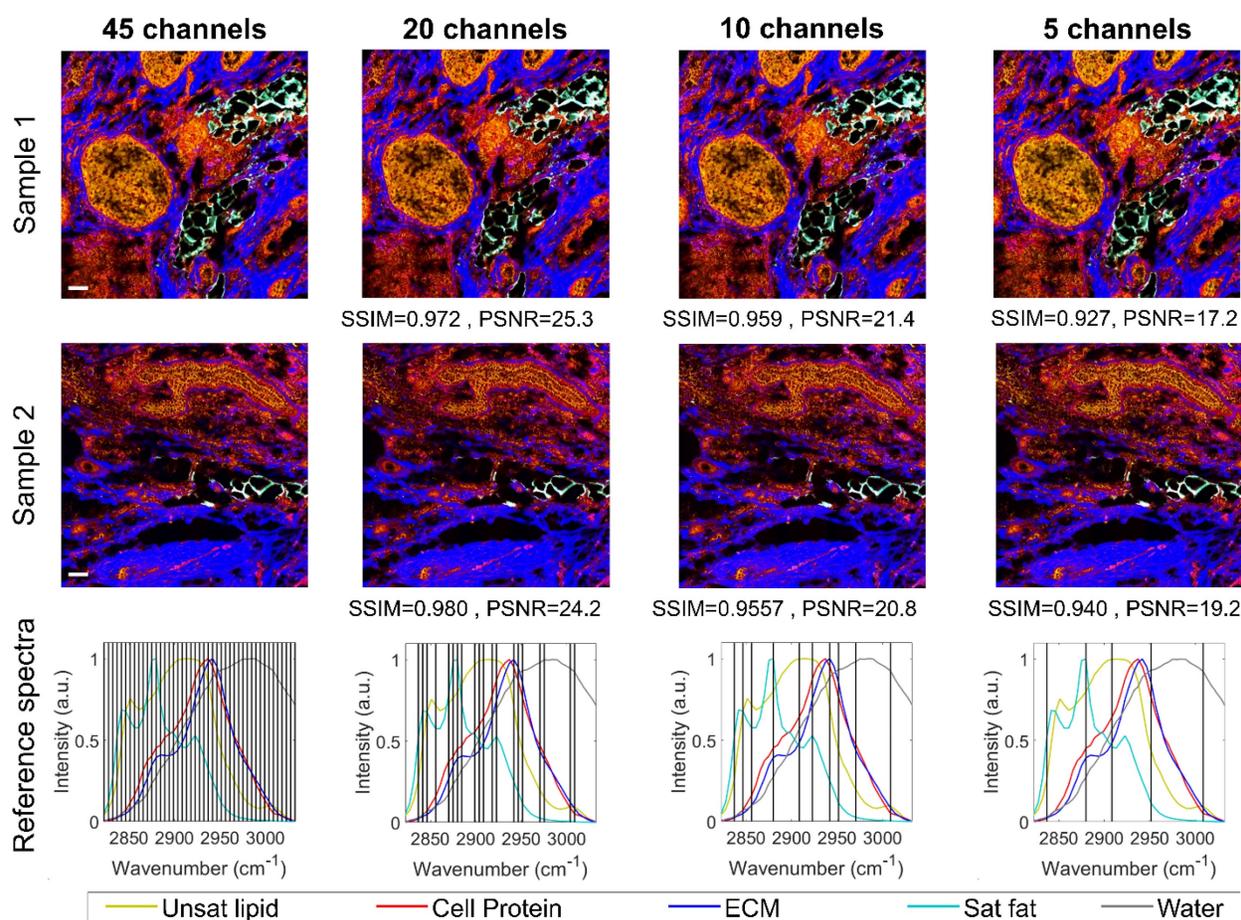


Figure 3. High-speed HC-SRH via spectral selective sampling. Scale bar: 50 μm . SSIM: structural similarity. PSNR: peak signal-to-noise ratio.

two-layer structure of the duct is still visible. ROI 7 is a blood vessel structure where the elastin around the vessel wall is shown as a mixture of ECM and cellular protein in HC-SRH. ROI 8 shows more intact fat cells close to the cancerous area. ROI 9 presents an ILC-like characteristic where cancer cells are aligned in a line shape (**Figure 4L**). ROI 10 shows a post-chemotherapy lesion consisting of newly grown collagen. The newly grown collagen is shown as a mixture of ECM and cellular protein in the HC-SRH, likely because the newly synthesized collagen is at an intermediate between fully grown collagen and cellular protein biomolecules. In summary, **Figure 4** indicates that HC-SRH has excellent contrast for duct, necrosis, stroma, blood vessel, fat cell, chemotherapy lesion, and epithelial cell morphology, which corresponds well with the H&E result. In addition, HC-SRH overperforms H&E in fat cell identification and distinguishing changes in the ECM composition.

HC-SRH with a compact FOPO laser

The results in **Figures 2-4** were acquired with a state-of-the-art SRS system operated with a solid-state OPO. The solid-state OPO is bulky and environment-sensitive, which is not suitable for clinical use. Furthermore, the complex optical

alignment, e.g., spectral focusing, to achieve hyperspectral scanning requires well-trained users. A rapid-tuning picosecond fiber laser facilitates the system operation and is suitable for clinical translation. We performed HC-SRH on a recently built FOPO-laser-based SRS system [29]. **Figure 5** compares the HC-SRH result acquired with the FOPO-laser-based system and that with the solid-state OPO-based system. Comparing **Figure 5A** and **5B**, we saw that selective spectral sampling can still preserve most of the HC-SRH quality using the same selected wavenumbers as in section 3.2. Equivalent image quality was achieved with fiber laser-based system as with solid-state OPO-based system (**Figure 5B** and **5C**).

SRS imaging of breast cancer tissue in the fingerprint window

Besides being clinically compatible, the FOPO-laser-based SRS system can complete fast tuning between two arbitrary wavenumbers within 20 milliseconds. This rapid spectral tuning capacity enables convenient spectroscopic SRS imaging across the whole biologically relevant Raman window. Fingerprint and C-H SRS hyperspectral images of a breast cancer tissue sample are shown in **Figure 6**.

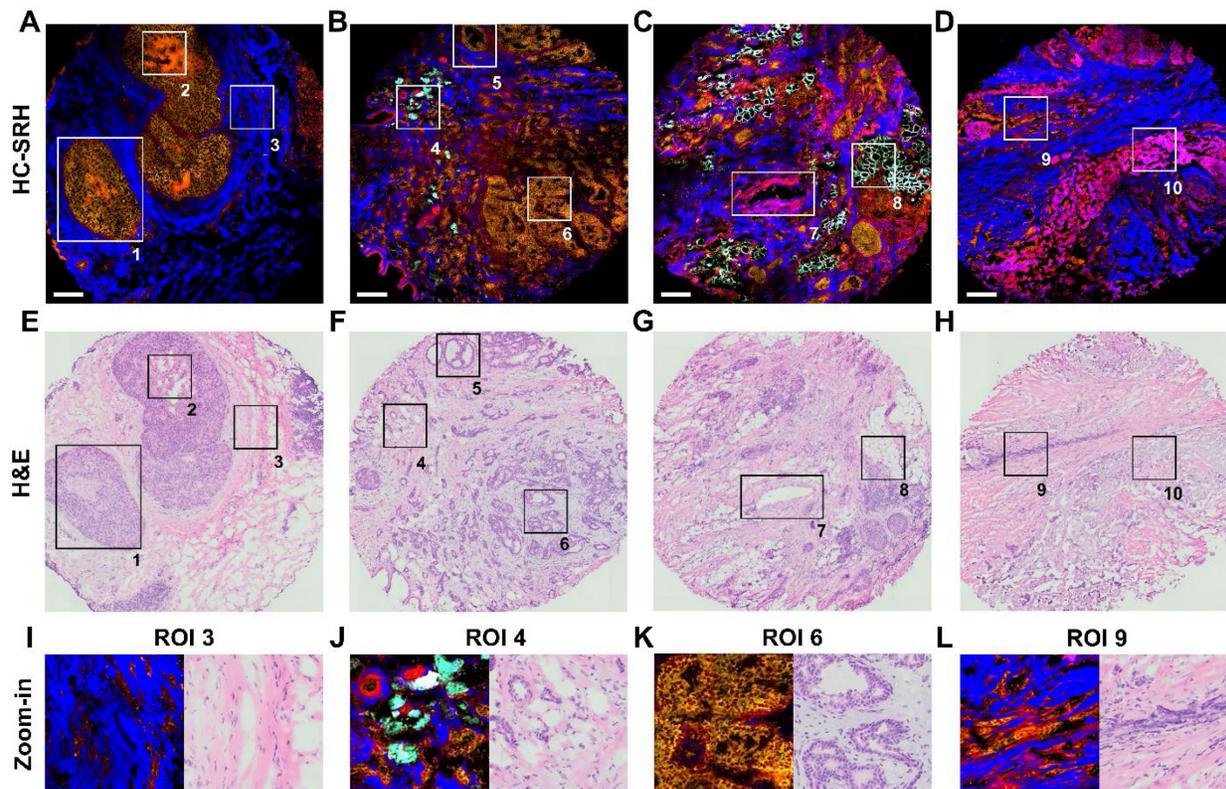


Figure 4. Breast tissue features revealed by HC-SRH. (A-D) HC-SRH of breast lesion tissue with 5-channel selective spectral sampling. A: typical ductal carcinoma in-situ (DCIS); B: mixture of usual ductal hyperplasia (UDH), normal duct, and invasive ductal carcinoma (IDC); C: IDC and invasive lobular carcinoma mixture (ILC); D: Chemotherapy lesion and some ILC residue. (E-H) H&E of neighboring tissue section corresponding to (A-D). (I-L) Zoom-in of ROIs 3,4,6,9. Color representation: Yellow: Unsaturated lipid, Red: Cell protein, Blue: Extracellular matrix, Cyan: Saturated fat. Scale bar: 200 μ m.

With the C-H HC-SRH (Figure 6A), we distinguished between different tissue components as introduced in section 3.1. The fingerprint images at the 6 characteristic Raman peaks are presented in Figure 6B. The 790 cm^{-1} and 1495 cm^{-1} images corresponding to nucleic acid peaks can clearly visualize the cancer cell nuclei, which is beyond the reach of C-H SRH. Phenylalanine and carotenoids have a characteristic peak at 1005 cm^{-1} . Images at this wavenumber can help reveal the detailed composition of protein and lipids. The image at 1250 cm^{-1} corresponding to the collagen further validates our ECM mapping in the C-H HC-SRH. Combining the 1300 cm^{-1} image for lipids and the 1745 cm^{-1} image for ester, we find that the lipids in the fat cell residue are highly esterized, but the lipids in cancer cell cytoplasm are not. Thus, combining information from the fingerprint region data and the C-H region data can provide valuable

information for more precise cancer diagnosis.

The fingerprint spectra of the 4 ROIs in Figure 6A are shown in Figure 6C. The 4 ROIs are selected to represent cancer cell cytoplasm, cancer cell nuclei, ECM, and fat cells. As expected, the cytoplasm spectrum shows a series of characteristic peaks for lipids and protein; the cell nuclei spectrum has many peaks corresponding to the nucleic acid; ECM spectra show many collagen features, and fat cell residue spectra are dominated by lipid characteristics. Another point to be noted is that we see a split ester peak in 1730 and 1745 cm^{-1} in the fat cell spectrum. This characteristic corresponds to a feature of solid-state fat [40], which is because the solid-state fat is better preserved than other lipids after tissue slicing. A detailed peak registration can be found in the supplementary material Table S1.

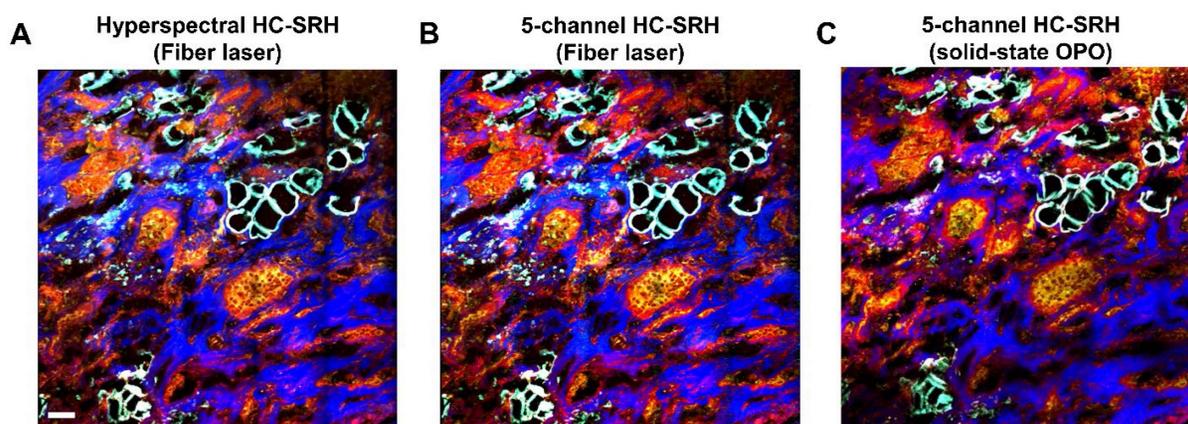


Figure 5. HC-SRH with a compact fiber laser. (A.) Hyperspectral HC-SRH (61 spectral channel) acquired by the FOPO-laser-based SRS system. (B.) Selective sampled HC-SRH (5 spectral channel) acquired by the FOPO-laser-based SRS system. The spectral channel position is the same with Figure 3. (C.) Selectively sampled HC-SRH (5 spectral channel) acquired in solid state OPO-based SRS system. Color representation: Yellow: Unsaturated lipid, Red: Cell protein, Blue: Extracellular matrix, Cyan: Saturated fat. Scale bar: $50\ \mu\text{m}$.

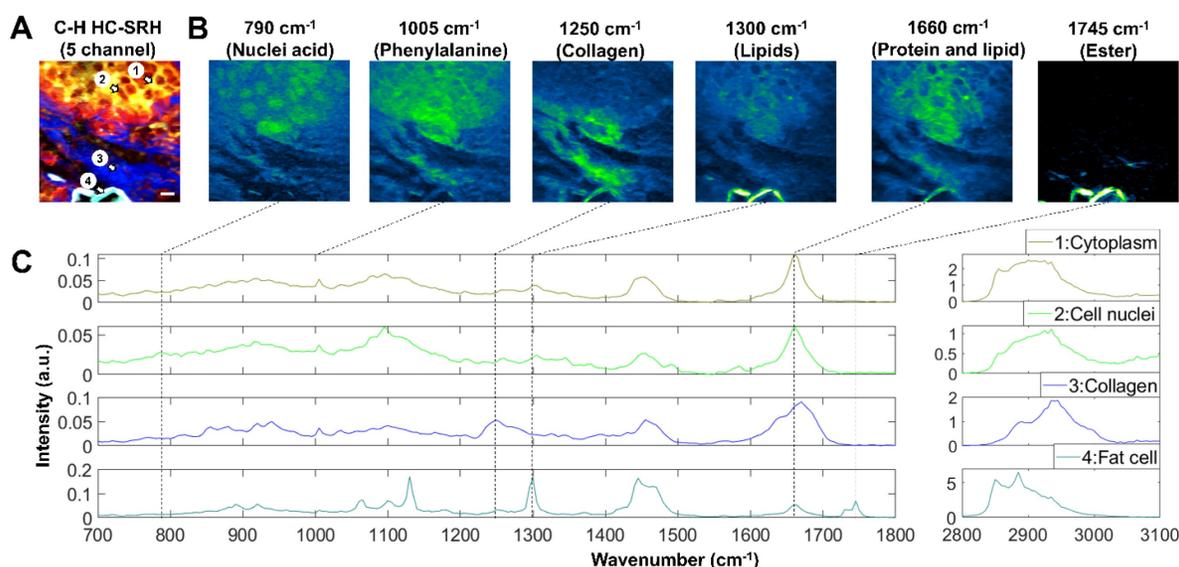


Figure 6. Fingerprint SRS imaging of breast cancer tissue. (A) C-H HC-SRH result. Color representation: Yellow: Unsaturated lipid, Red: Cell protein, Blue: Extracellular matrix, Cyan: Saturated fat. (B) Fingerprint SRS result at representative spectral position in the same FOV of (A). (C) Multi-window SRS spectra of the 4 selected ROIs in (A). Scale bar: $10\ \mu\text{m}$.

Discussion

In this study, we have developed stain-free HC-SRH that provides breast cancer histology with rich chemical information. Compared to the well-developed two-color SRH, HC-SRH provides extra chemical information that can benefit cancer diagnosis in two aspects. First, HC-SRH has the potential to improve cancer classification accuracy which is of great importance in clinics. Machine learning has been widely applied to histology results to obtain diagnostic information automatically. Although decent classification accuracy can be obtained from tissue morphology, studies have indicated that chemical information can further improve classification performance [41]. For clinical application, classification accuracy is of great importance, and every improvement is worthwhile. In addition, chemical information gained from HC-SRH can contribute greatly to our knowledge about cancer biology which can accelerate the discovery of new biomarkers. The informative fingerprint window is especially promising for identifying cancer biomarkers [5]. For example, microcalcifications [17, 18] and carotenoids [42] are indicated as biomarkers for breast cancer development, which have characteristic peaks in the fingerprint window. Other than relying on the morphological-based machine learning classification, HC-SRH can deepen our understanding of subtle chemical changes involved in cancer progression.

In order to further advance the HC-SRH toward clinical applications, several future works must be carried out. Firstly, an HC-SRH classification model based on either machine learning or specific chemical biomarkers needs to be established to generate diagnostic information, which can help the HC-SRH to be used by the doctor and fit into the existing clinical workflow. Such a model is missing in this work due to the limited number of samples. Moreover, HC-SRH would better be performed on fresh, unsliced biopsy samples, which would minimize the sample preparation requirement and avoid artifacts caused by the tissue slicing. One slicing artifact we observed is the loss of fat cells, resulting in dark holes in our imaging results. Additionally, the water channel of HC-SRH could have been more informative because the original water content in the tissue is altered during the sample preparation, i.e., frozen sectioning and external PBS. The feasibility of applying HC-SRH on fresh unsliced biopsy can be shown by the two-color SRH work performed on fresh breast needle biopsy samples [15]. As HC-SRH has the same imaging depth as two-color SRH, HC-SRH has the promise to be applied to unsliced

biopsy to get the image of the tissue surface and performed on the fresh samples.

Conclusion

We have developed a new approach for label-free cancer histology called high-content stimulated Raman histology (HC-SRH). The HC-SRH in the C-H window allows the mapping of major chemical contents in the breast tissue and provides excellent contrast for various tissue components. Through implementing spectral selective sampling, we boosted the speed of HC-SRH by ~1 order without sacrificing image quality. We also demonstrated that the HC-SRH is robust with a clinical-compatible fiber laser system. With the rapid, widely tuning capability of the FOPO, we extended the spectral coverage of the HC-SRH to the fingerprint window, which provides extra contrast for nucleic acid, amino acid, and solid-state fat in breast tissue. Collectively, these results show that HC-SRH is a promising tool for label-free cancer histology with rich chemical information.

Abbreviations

HC-SRH: high-content stimulated Raman histology; CARS: coherent anti-Stokes Raman scattering; SRS: stimulated Raman scattering; SRH: stimulated Raman histology; H&E: hematoxylin and eosin; ECM: extracellular matrix; LASSO: least absolute shrinkage and selection operator; OPO: optical parametric oscillator; AOM: acousto-optic modulator; NA: numerical aperture; FOV: field-of-view; SHG: second harmonic generation; FOPO: fiber optical parametric oscillator; PBS: phosphate-buffered saline; RFE: recursive feature elimination; MSE: mean square error; SSIM: structural similarity; PSNR: peak signal-to-noise ratio; DCIS: ductal carcinoma in-situ; UDH: usual ductal hyperplasia; IDC: invasive ductal carcinoma; ILC: invasive lobular carcinoma; ROI: regions of interest.

Supplementary Material

Supplementary figures and table.

<https://www.thno.org/v14p1361s1.pdf>

Acknowledgements

This project is supported by funds provided by Hologic to Boston University. We thank Dr. Linda Han for the helpful discussion on breast cancer diagnosis.

Competing Interests

The authors have declared that no competing interest exists.

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021; 71: 209-49.
- Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, et al. *Breast Cancer Statistics, 2022.* *CA Cancer J Clin.* 2022; 72: 524-41.
- van Dooyeweert C, van Diest PJ, Willems SM, Kuijpers C, van der Wall E, Overbeek LIH, et al. Significant inter- and intra-laboratory variation in grading of invasive breast cancer: A nationwide study of 33,043 patients in the Netherlands. *Int J Cancer.* 2020; 146: 769-80.
- Holzer I, Farr A, Tan Y, Deutschmann C, Leser C, Singer CF. Receptor Discordance of Metastatic Breast Cancer Depending on the Molecular Subtype. *Breast Care (Basel).* 2020; 15: 648-54.
- Lazaro-Pacheco D, Shaaban AM, Rehman S, Rehman I. Raman spectroscopy of breast cancer. *Applied Spectroscopy Reviews.* 2020; 55: 439-75.
- Cui S, Zhang S, Yue S. Raman Spectroscopy and Imaging for Cancer Diagnosis. *J Healthc Eng.* 2018; 2018: 8619342.
- Cicerone MT, Camp CH. Histological coherent Raman imaging: a prognostic review. *Analyst.* 2017; 143: 33-59.
- Camp CH, Jr., Lee YJ, Heddeston JM, Hartshorn CM, Hight Walker AR, Rich JN, et al. High-Speed Coherent Raman Fingerprint Imaging of Biological Tissues. *Nat Photonics.* 2014; 8: 627-34.
- Zhang D, Wang P, Slipchenko MN, Cheng JX. Fast vibrational imaging of single cells and tissues by stimulated Raman scattering microscopy. *Acc Chem Res.* 2014; 47: 2282-90.
- Ji M, Orringer DA, Freudiger CW, Ramkissoon S, Liu X, Lau D, et al. Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy. *Sci Transl Med.* 2013; 5: 201ra119.
- Ji M, Lewis S, Camelo-Piragua S, Ramkissoon SH, Snuderl M, Venneti S, et al. Detection of human brain tumor infiltration with quantitative stimulated Raman scattering microscopy. *Science Translational Medicine.* 2015; 7: 309ra163.
- Orringer DA, Pandian B, Niknafs YS, Hollon TC, Boyle J, Lewis S, et al. Rapid intraoperative histology of unprocessed surgical specimens via fibre-laser-based stimulated Raman scattering microscopy. *Nat Biomed Eng.* 2017; 1.
- Hollon TC, Pandian B, Adapa AR, Urias E, Save AV, Khalsa SSS, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat Med.* 2020; 26: 52-8.
- Liu Z, Su W, Ao J, Wang M, Jiang Q, He J, et al. Instant diagnosis of gastroscopic biopsy via deep-learned single-shot femtosecond stimulated Raman histology. *Nat Commun.* 2022; 13: 4050.
- Yang Y, Liu Z, Huang J, Sun X, Ao J, Zheng B, et al. Histological diagnosis of unprocessed breast core-needle biopsy via stimulated Raman scattering microscopy and multi-instance learning. *Theranostics.* 2023; 13: 1342-54.
- Ao J, Shao X, Liu Z, Liu Q, Xia J, Shi Y, et al. Stimulated Raman Scattering Microscopy Enables Gleason Scoring of Prostate Core Needle Biopsy by a Convolutional Neural Network. *Cancer Res.* 2023; 83: 641-51.
- Shin KS, Laohajaratsang M, Men S, Figueroa B, Dintzis SM, Fu D. Quantitative chemical imaging of breast calcifications in association with neoplastic processes. *Theranostics.* 2020; 10: 5865-78.
- Yang Y, Yang Y, Liu Z, Guo L, Li S, Sun X, et al. Microcalcification-Based Tumor Malignancy Evaluation in Fresh Breast Biopsies with Hyperspectral Stimulated Raman Scattering. *Anal Chem.* 2021; 93: 6223-31.
- Fu D, Holtom G, Freudiger C, Zhang X, Xie XS. Hyperspectral imaging with stimulated Raman scattering by chirped femtosecond lasers. *J Phys Chem B.* 2013; 117: 4634-40.
- Laptenok SP, Rajamanickam VP, Genchi L, Monfort T, Lee Y, Patel, II, et al. Fingerprint-to-CH stretch continuously tunable high spectral resolution stimulated Raman scattering microscope. *J Biophotonics.* 2019; 12: e201900028.
- Talari ACS, Rehman S, Rehman IU. Advancing cancer diagnostics with artificial intelligence and spectroscopy: identifying chemical changes associated with breast cancer. *Expert Rev Mol Diagn.* 2019; 19: 929-40.
- Melitto AS, Arias VEA, Shida JY, Gebrim LH, Silveira L, Jr. Diagnosing molecular subtypes of breast cancer by means of Raman spectroscopy. *Lasers Surg Med.* 2022; 54: 1143-56.
- Haka AS, Shafer-Peltier KE, Fitzmaurice M, Crowe J, Dasari RR, Feld MS. Diagnosing breast cancer by using Raman spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America.* 2005; 102: 12371.
- de Andrade Natal R, Adur J, Cesar CL, Vassallo J. Tumor extracellular matrix: lessons from the second-harmonic generation microscopy. *Surgical and Experimental Pathology.* 2021; 4.
- Abramczyk H, Brozek-Pluska B. New look inside human breast ducts with Raman imaging. Raman candidates as diagnostic markers for breast cancer prognosis: Mammaglobin, palmitic acid and sphingomyelin. *Anal Chim Acta.* 2016; 909: 91-100.
- Marques MPM, Batista de Carvalho ALM, Mamede AP, Dopplapudi A, Garcia Sakai V, Batista de Carvalho LAE. Role of intracellular water in the normal-to-cancer transition in human cells-insights from quasi-elastic neutron scattering. *Struct Dyn.* 2020; 7: 054701.
- Lin H, Lee HJ, Tague N, Lugagne J-B, Zong C, Deng F, et al. Microsecond fingerprint stimulated Raman spectroscopic imaging by ultrafast tuning and spatial-spectral learning. *Nature Communications.* 2021; 12.
- Tan Y, Lin H, Cheng J-X. Profiling single cancer cell metabolism via high-content SRS imaging with chemical sparsity. *Science Advances.* 2023; 9: eadg6061.
- Ni H, Lin P, Zhu Y, Zhang M, Tan Y, Zhan Y, et al. Multiwindow SRS Imaging Using a Rapid Widely Tunable Fiber Laser. *Analytical Chemistry.* 2021; 93: 15703-11.
- Brinkmann M, Fast A, Hellwig T, Pence I, Evans CL, Fallnich C. Portable all-fiber dual-output widely tunable light source for coherent Raman imaging. *Biomed Opt Express.* 2019; 10: 4437-49.
- Liao CS, Choi JH, Zhang D, Chan SH, Cheng JX. Denoising Stimulated Raman Spectroscopic Images by Total Variation Minimization. *J Phys Chem C Nanomater Interfaces.* 2015; 119: 19397-403.
- Wright AM, Howard AA, Howard JC, Tschumper GS, Hammer NI. Charge transfer and blue shifting of vibrational frequencies in a hydrogen bond acceptor. *J Phys Chem A.* 2013; 117: 5435-46.
- Maiti NC, Apetri MM, Zagorski MG, Carey PR, Anderson VE. Raman Spectroscopic Characterization of Secondary Structure in Natively Unfolded Proteins: α -Synuclein. *Journal of the American Chemical Society.* 2004; 126: 2399-408.
- Chajes V, Niyongabo T, Lanson M, Fignon A, Couet C, Bougnoux P. Fatty-acid composition of breast and iliac adipose tissue in breast-cancer patients. *Int J Cancer.* 1992; 50: 405-8.
- Nieva C, Marro M, Santana-Codina N, Rao S, Petrov D, Sierra A. The lipid phenotype of breast cancer cells characterized by Raman microspectroscopy: towards a stratification of malignancy. *PLoS One.* 2012; 7: e46456.
- Ranjit S, Dvornikov A, Stakic M, Hong SH, Levi M, Evans RM, et al. Imaging Fibrosis and Separating Collagens using Second Harmonic Generation and Phasor Approach to Fluorescence Lifetime Imaging. *Sci Rep.* 2015; 5: 13378.
- Le TT, Langohr IM, Locker MJ, Sturek M, Cheng JX. Label-free molecular imaging of atherosclerotic lesions using multimodal nonlinear optical microscopy. *J Biomed Opt.* 2007; 12: 054007.
- Vaidya Y, Vaidya P, Vaidya T. Ductal Carcinoma In Situ of the Breast. *Indian J Surg.* 2015; 77: 141-6.
- Bodis S, Siziopikou KP, Schnitt SJ, Harris JR, Fisher DE. Extensive apoptosis in ductal carcinoma in situ of the breast. *Cancer.* 1996; 77: 1831-5.
- Czamará K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M. Raman spectroscopy of lipids: a review. *Journal of Raman Spectroscopy.* 2015; 46: 4-20.
- Manifold B, Men S, Hu R, Fu D. A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nature Machine Intelligence.* 2021; 3: 306-15.
- Eliassen AH, Liao X, Rosner B, Tamimi RM, Tworoger SS, Hankinson SE. Plasma carotenoids and risk of breast cancer over 20 y of follow-up. *Am J Clin Nutr.* 2015; 101: 1197-205.